

Министерство образования и науки Российской Федерации
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г.ЧЕРНЫШЕВСКОГО»

Кафедра информатики и программирования

**РАСПОЗНАВАНИЕ СИМВОЛОВ С ПОМОЩЬЮ НЕЙРОННЫХ
СЕТЕЙ**

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студента 4 курса 441 группы
направления Математическое обеспечение и администрирование
информационных систем
факультета компьютерных наук и информационных технологий
Добрынина Артема Михайловича

Научный руководитель

к.э.н., доцент

Л.В. Кабанова

Зав. кафедрой

к.ф.-м.н.

М.В. Огнева

Саратов 2018

ВВЕДЕНИЕ

Актуальность темы. Работа с текстом встречается в жизни человека повсеместно: прочитать почту ранним утром, узнать о проведении ремонтных работ в доме по дороге на работу, определить нужный номер автобуса – всё это лишь малая часть того, с чем сталкивается каждый.

Анализ текста является элементарной функцией головного мозга. И довольно комплексной в компьютерных технологиях. Для каждой сферы деятельности человека требуется свой анализатор: определить номер автомобиля нарушителя, проверить индекс почтового отделения, получить лист печатного текста по фотографии – для всего этого нужен собственный анализатор текста.

Вышесказанное определило *цель бакалаврской работы*: реализовать нейронную сеть, способную распознавать печатный текст на фотографии и переводить его в текст.

Исходя из этого, были поставлены следующие *задачи*:

- Исследовать методы программного распознавания символов;
- Изучить эффективность использования нейронных сетей при распознавании символов;
- Изучить методы оптимизации анализа текста с помощью нейронных сетей;
- Реализовать нейронную сеть, способную распознавать символы;
- Проанализировать временную сложность работы алгоритма.

Методологические основы распознавания символов с помощью нейронных сетей представлены в работах Я. Голдфеллоу[3], И. А. Багровой[8], Я. Лекуна[13, 14, 15, 17].

Практическая значимость бакалаврской работы. В ходе выполнения практической части была реализована система распознавания символов на изображении. Данное приложение может быть использовано для ускорения набора печатного текста.

Структура и объём работы. Бакалаврская работа состоит из введения, 3 разделов, заключения, списка использованных источников и 2 приложений. Общий объем работы – 49 страниц, из них 46 страниц – основное содержание, включая 16 рисунков и 3 таблицы, цифровой носитель в качестве приложения, список использованных источников информации – 20 наименований.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Первый раздел «Теоретические основы анализа текста» посвящён различным методам обработки и анализа текста. Были рассмотрены методы распознавания образов, а также особенности использования нейронных сетей[1, 2, 3, 4].

В подразделе «Общие сведения о распознавании образов» были подробно рассмотрены методы распознавания образов. Они представляют из себя большой класс различных подходов, которые решают разнообразные задачи. Из всего множества данных методов были выделены три группы и подробно рассмотрены относящиеся к ним методы. Группы подразделены на:

- предварительную фильтрацию и подготовку изображения;
- логическую обработку результатов фильтрации;
- алгоритмы принятия решений на основе логической обработки.

В подразделе «Применение методов распознавания образов для символов» были рассмотрены теоретические основы и архитектура свёрточных нейронных сетей.

Второй раздел «Методы распознавания текста» посвящён подробному анализу методов распознавания текста и их практической значимости в системе распознавания образов.

В подразделе «Бинаризация» были определены наиболее эффективные алгоритмы бинаризации, выделены и проанализированы преимущества и недостатки каждого алгоритма[5].

В подразделе «Классификатор» были рассмотрены различные типы классификаторов, построенные на основе различных алгоритмов. В частности, были рассмотрены алгоритмы на основе теоремы Байеса, признаков Хаара и методе Виолы-Джонса[6].

В подразделе «Нейронная сеть» были даны базовые определения для нейронных сетей, подробно рассмотрены методы обучения сети, а также различные виды нейронных сетей. В частности, были рассмотрены

многослойные перцептроны, свёрточные нейронные сети и ансамбли синапсов[7, 8, 9, 10, 11, 12].

Третий раздел «Реализация системы распознавания текста» посвящён программной реализации нейронной сети на языке Java с применением технологий Maven, JavaCV и Java Concurrency. В качестве среды разработки было выбрано решение IntelliJ IDEA от JetBrains.

JavaCV является реализацией библиотеки OpenCV, изначально написанной для C/C++. Она предоставляет удобный инструментарий для обработки фото и видео.

В подразделе «Постановка задачи и реализуемые методы» подробно описаны требования к поставленной задаче и выбранные методы для её решения. Также рассмотрена и проанализирована топология используемой свёрточной нейронной сети[13].

В подразделе «Разработка алгоритма» описаны все шаги алгоритма распознавания символов. Сначала входные данные поступают на вход алгоритма сегментации, после чего получается тестовая выборка элементов, пригодных для обработки нейросетью. Был подробно описан алгоритм сегментации, а также выявлены наиболее подходящие приёмы обработки изображения. Были получены и проанализированы промежуточные результаты сегментации, что поспособствовало избежать крупных ошибок на ранних этапах обработки. Также был подробно рассмотрен процесс формирования, обучения и тестирования нейросети. Кроме этого, был подробно рассмотрен алгоритм обратного распространения ошибки, что является очень важным для свёрточной нейронной сети. В частности, были рассмотрены различные варианты расчёта ошибки для различных типов слоёв сети[3, 9, 13, 14, 15, 16, 17, 18, 19, 20].

В подразделе «Результаты работы программы и временная сложность» были представлены результаты работы нейронной сети на тестовых значениях. Также была выявлена разница точности распознавания символов для разных обучающих выборок. Полученная нейронная сеть, обученная на

наиболее эффективном наборе данных, имеет точность распознавания обучающей выборки 97,7% и тестовой выборки 78,5%. В связи с этим выявлены особенности данной системы распознавания символов и влияние алгоритма сегментации на окончательный результат процента распознавания. Также была проведена оценка временной сложности алгоритма. Запуск алгоритма происходил в многопоточном режиме. По результатам данной оценки, время работы алгоритма с обученной нейронной сетью составляет 5 секунд, что является хорошим показателем быстродействия системы.

ЗАКЛЮЧЕНИЕ

В ходе данной работы были выполнены все поставленные задачи, что позволило достичь заявленной цели – реализовать нейронную сеть, способную распознавать печатный текст на фотографии и переводить его в текст. Данная нейронная сеть является свёрточной, что обеспечивает большую точность распознавания.

В ходе выполнения задания был реализован алгоритм сегментации изображения с помощью библиотеки OpenCV. Это помогло решить задачу определения символов на изображении.

Также были исследованы методы программного распознавания символов. Они заключаются в комбинации различных методов обработки изображения и нахождения признаков символов, а также анализа и классификации этих признаков.

Были изучены эффективность использования нейросетей и методы оптимизации анализа текста с помощью нейронных сетей. Для эффективного распознавания система анализа текста может содержать несколько нейросетей для нахождения признаков и классификации.

Была реализована нейросеть, способная распознавать символы. Её результат распознавания на обучающих выборках – 97,7%, на тестовых – 78,5%. Ошибки, кроме прочего, обусловлены ошибками сегментации. Кроме того, в тестах были использованы шрифты, не участвующие в обучающих выборках.

Временная сложность данного алгоритма составляет примерно 5 секунд на обученной нейросети. Большую часть этого времени занимает сегментация. Кроме того, алгоритм адаптирован под многопоточную работу, что позволяет обеспечить лучшее время работы на мощных процессорах.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. *Гонсалес Р., Вудс Р.* Цифровая обработка изображений (3-е издание, исправленное и дополненное) — М.: Техносфера, 2012.
2. *Яне Б.* Цифровая обработка изображений (Пер. с англ. А.М. Измайловой) — М.: Техносфера, 2007.
3. Обработка цветных изображений – Техническое зрение [Электронный ресурс] URL:
http://wiki.technicalvision.ru/index.php/Обработка_цветных_изображений
(Дата обращения: 08.05.18)
4. *Ian J. Goodfellow, Yaroslav Bulatov, Julian Ibarz, Sacha Arnoud, Vinay Shet* Multi-digit Number Recognition from Street View Imagery using Deep Convolutional Neural Networks — Google Inc, 2014
5. *Y. LeCun, Y. Bengio* Convolutional Networks for Images, Speech, and Time-Series, in *Arbib, M. A. (Eds), The Handbook of Brain Theory and Neural Networks*, MIT Press, 1995
6. *Арлазаров В.Л., Куратов П.А., Славин О.А.* Распознавание строк печатных текстов // Сб. трудов ИСА РАН «Методы и средства работы с документами». — М.: Эдиториал УРСС, 2000.
7. *Круг П.Г.* Нейронные сети и нейрокомпьютеры: Учебное пособие по курсу «Микропроцессоры». – М.: Издательство МЭИ, 2002.
8. *R. Marc'Aurelio, C. Poultney, S. Chopra, Y. LeCun:* Efficient Learning of Sparse Representations with an Energy-Based Model, in *J. Platt et al. (Eds), Advances in Neural Information Processing Systems (NIPS 2006)*, MIT Press, 2006
9. *J. Bouvrie.* Notes on Convolutional Neural Networks, MIT Press, 2006.
10. *Багрова И. А., Грицай А. А., Сорокин С. В., Пономарев С. А., Сытник Д. А.* Выбор признаков для распознавания печатных кириллических символов // Вестник Тверского Государственного Университета 2010 г.
11. *Панфилов С. А.* Методы и программный комплекс моделирования алгоритмов управления нелинейными динамическими системами на

- основе мягких вычислений. Диссертация на соискание ученой степени кандидата технических наук. Тверь, 2005.
12. *Berber M.* Three robust features extraction approaches for facial gender classification. *Visual Comput*, 2013.
 13. Проект СПбГУ Открытый код: распознавание текстовых изображений [Электронный ресурс] URL: <http://ocr.apmath.spbu.ru> (Дата обращения: 08.05.2018)
 14. *Квасников В.П., Дзюбаненко А.В.* Улучшение визуального качества цифрового изображения путем поэлементного преобразования // *Авиационно-космическая техника и технология.* – 2009
 15. *Melin P., Urias J., Solano D., Soto M., Lopez M., Castillo O.* Voice Recognition with Neural Networks, Type-2 Fuzzy Logic and Genetic Algorithms. *Engineering Letters*, 13:2, 2006.
 16. *Y. LeCun, J. S. Denker, S. Solla, R. E. Howard, L. D. Jackel* Optimal Brain Damage, in Touretzky, David (Eds), *Advances in Neural Information Processing Systems 2 (NIPS*89)*, Morgan Kaufman, Denver, CO, 1990
 17. *S. S. Liew.* Gender classification: A convolutional neural network approach, 2016.
 18. *D. H. Hubel, T. N. Wiesel* Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex, 1961.
 19. *Y. LeCun, L. Bottou, G. Orr, K. Muller* Efficient BackProp, in Orr, G. and Muller K. (Eds), *Neural Networks: Tricks of the trade*, Springer, 1998
 20. *Y. LeCun.* Gradient-Based Learning Applied to Document Recognition, 1998.